TECHNOLOGY IN MATHEMATICS CLASSROOMS: A META-ANALYSIS OF THE RECENT LITERATURE

By

QING LI*

XIN MA**

ABSTRACT

The increasing use of technology in education does set off a flurry of research studies that focus on the successfulness and effectiveness of technology in elementary and secondary education. In this study, a comprehensive meta-analysis of research literature on technology in relation to mathematics teaching and learning was conducted. Particularly, the following main research questions were addressed in this meta-analysis:

1. What is the magnitude of the effects of technology on schooling outcomes concerning mathematics education?

2. How does the magnitude of the effects of technology fluctuate in response to various study features (e.g., gender, age, race) and design features (e.g., randomization, sample size, instruments)?

Based on a total of 81 independent findings extracted from 39 studies involving a total of 59,147 learners, the results of the series of meta-analysis conducted in this review indicate that technology can affect mathematics teaching and learning.

Keywords: meta-analysis, technology, math learning, achievement, attitude, behavior.

INTRODUCTION

Educational technology, as the term is used in the realm of education, refers to the technical means thot are used to support teaching and leorning, such as computers, calculators, educational software programs, interactive media, and tele-communication systems. The use of educationol technology hos become increasingly popular in elementary and secondary schools over the past several decades. There is little doubt that technology has become a ubiquitous tool for teoching ond learning. The National Council of Teachers of Mathematics (NCTM, 2000) emphasizes the importance of the use of technology in mathematics education, stating that "technology is essential in teaching and learning mathematics; it influences the mathematics that is taught and enhances students' learning" (p. 2).

Although technology has great potential to impact the teaching and learning of mathematics, the presence of technology does not warront desirable schooling outcomes concerning mathematics education (Clark, 1983; Li, 2004). Successful and effective use of technology for the teoching and learning of mathematics depends upon sound teoching and learning strategies that come from a thorough understonding of the effects of technology on mathematics education (Albright & Graf, 1992; Coley, Cradleer, & Engel, 2000).

The increasingly popular use of technology in education has resulted in a flurry of research studies (referred as primary studies) that focus on the successfulness ond effectiveness of technology in elementary and secondary education. How technology can be used successfully and effectively to affect the teaching ond learning of mathematics in K-12 classrooms is the key research question that many primary studies have attempted to address. Unsurprisingly, findings have not been consistent, especially when technology use is compounded with other factors such as student characteristics (e.g., gender, ability) (Royer, Greene, & Anzalone, 1994; Salerno, 1995), student group composition (Brush, 1997; Xin, 1999), and teaching methods (Farrell, 1996; Hecht, Roberts, Schoon, & Fansler, 1995; Shyu, 1999).

As research evidence accumulates on this educational issue, reseorch synthesis becomes necessary to make sense from a large body of research literature. In this study, a comprehensive meta-analysis of research literature on technology in relation to mathematics teaching and

learning wos conducted. The need for the present meto-analysis is two-fold. First, although several review studies are available camparing technalagy with traditianal instructional approaches (Kulik, 2003; Kulik, Schwalb, & Kulik, 1982; Parr, 2003), there has been no focus on the effects af technalagy an teaching and learning mathematics in schaal, as ane af the care academic subjects. Second, primary studies have provided inconsistent results concerning the effects of technology in mathematics classraams, calling "far a systematic integration of the literature both for theory development and for pedagogical guidance" (Lou, Abromi, & d'Apallonia, 2001, p. 451). This study, therefore, systematically reviews the existing literature focusing on the following main reseorch questions:

1. What is the magnitude of the effects of technology on schooling outcomes concerning mathematics educatian?

2. How does the magnitude af the effects of technology fluctuate in respanse ta variaus study features (e.g., gender, age, race) and design features (e.g., randomization, sample size, instruments)?

Review of Related Literature

This section briefly reviews the research related to educational use of technology in K-12 mathematics teaching and learning. This review nat anly provides the background informatian for this study, but also helps us to identify features to consider in our quantitotive integrotion of the effects af technalagy in mathematics education.

Type of Technology

Different types of technology and a variety of computer programs have been developed and used in an attempt ta enhance mathematics teaching and learning in the post several decades. These include Logo, Spreodsheet, the Internet, and various calculators/graphing calculatars. "Guided by different learning theories, philosophies, or developments in technology, each type of technology oppeors to have distinct characteristics, purpases and different ways ta facilitate student learning" (Lau et al., 2001, p. 452). One main categary af technology integration in mothematics classrooms is the use of software specifically designed for mathematics leorning. Abundant softwares are avoilable in the market and examples af such programs are Geameter's Sketchpad and Mathematica. Many teachers also like to use general-purpose technological tools in mathematics educatian. These taals include ward processing taals, spreadsheet, multimedia and hypermedia. Further, different communication tools such as email, computer-conferences, video-conferencing, and the Internet are used in mathematics classraams which enable communication and information sharing amongst geographically dispersed learners.

Extensive research has been conducted on the use of technology in mathematics education. In a meta-analysis af twenty seven studies, Christmann, Badgett, and Lucking (1997) compared the academic achievement of students, studying in grade 6 to 12, who received either traditianal instructian ar traditional instruction supplemented with Computer-Assisted Instruction (CAI) across eight curricular areos. For mothematics, o total of 15 of effect sizes were calculated with the mean effect size af 0.179. The researchers concluded that the effect of CAI on mathematics leorning was weok. It is important to note thot the studies included in their meta-analysis, particularly the anes involving mathematics, were either papers (journal articles or conference presentations) published before 1990 ar dissertations.

The "Adventures of Jasper Woodbury" mathematics program developed by a team at Vanderbilt University has been widely implemented in many areas of the U.S and around the world. Based on the theory of anchored instruction, the pragram uses videa and multimedia camputing technology ta pravide prablem-scenarias aiming to help students to develop necessary skills ond knowledge for problem solving and critical thinking. Implementatian af this pragram yielded same interesting findings (Mushi, 2000; Shyu, 2000). Far example, an evaluation (Mushi, 2000) of the use of this mathematics pragram in 8 schaals in Chicaga praduced twa important results. First, quantitative analysis af the data callected from 1,275 grade 5 to 8 students showed no significant

gains in student knowledge, skills, and attitudes. Qualitative data collected, however, indicated that learning mathematics thraugh media was interesting ta students and had made positive impact on students' attitudes towards mathematics.

Many teachers use computer software programs as supplements to the regular curriculum or as instructional alternatives. Funkhauser's study (1993) demanstrated that the use of a problem-solving computer software had contributed to a positive change in the attitudes of secandary students tawards mathematics. Significant gains in problem-salving ability and knawledge of math content were also observed.

In mathematics classrooms, particularly in elementary mathematics classrooms, manipulatives have been used intensively ta help building a faundatian for students to understand abstract concepts. The increasing access to computer technology in schools inevitably resulted in some enthusiasm far the use of virtual manipulatives for mathematics learning. Virtual manipulatives can be briefly defined as "replicas of physical manipulatives that can be accessed through the Internet". In addition, it is "an interactive, web-based visual representatian af a dynamic abject that presents appartunities far constructing mathematical knowledge" (Moyer, Bolyard, & Spikell, 2002, p. 372). One advantage af virtual manipulatives, accarding ta Reimer and Mayer (2005), is the capability to connect dynamic visual images with abstract symbols a limitation of regular manipulatives. Same researchers have described haw virtual manipulatives can be used ta teach fraction concepts for elementary students (Suh, Moyer, & Heo, 2005). Others have examined haw juniar high students used virtual pattern blacks, virtual platanic salids, and virtual geoboards to explore geometric concepts (Moyer & Bolyard, 2002). A variety of studies examined the virtual manipulative taal in mathematics classraams and cancluded that there exists a positive impact of such tools on student achievement and attitudes (Char, 1989; Kieran & Hillel, 1990; Thampson, 1992). When virtual manipulatives are used in cambinatian with regular manipulatives, researchers also found positive results (Ball, 1988; Terry, 1996). However, results from other studies indicated no significant gains in students' achievement (Kim, 1993; Nute, 1997). The results af research in this area are inconclusive and the amount of research on high-quality "virtual manipulatives is so limited that a judgment abaut their patential uses in mathematics instruction is entirely speculative" (Reimer & Mayer, 2005, p. 8).

Calculatars, including graphing calculatars, is anather tool that is used extensively in mathematics classrooms. Having been around for several decades, the functianalities af calculatars have cantinually been expanded. For example, the graphing calculatars include "numerical calculations, the graphing of functions, the manipulation of lists of data, and the calculatian and display af statistical graphs" (Janes, 2005, p. 31). Many research studies of graphing calculators focused on teaching and learning algebra in secandary schaals. Few explared the use af such taals in ather tapics like statistics.

Accarding ta Janes (2005), the general trend identified in research studies is that using graphing calculators can enable students to approach situations graphically, numerically and symbalically, and can suppart students' visualizatian, allawing them ta explare situatians which they may not otherwise be able to tackle (and thus perhaps enable them ta take their mathematics ta a mare advanced level). In this way, using graphing calculators can lead to higher achievement among students, perhaps through increased student use of graphical salutian strategies, impraved understanding af functions, and increased teacher time spent on presentation and explanation of graphs, tables, and prablem salving activities. (p. 31).

Learning conditions and learner characteristics

The literatures on the exploration of educational use of technalagy in mathematics classraams suggest that the effect af technalagy an learning may depend an the learning environment. Some researchers have attributed students' academic successes and attitudinal changes ta methads fram a pedagagical refarm, rather than merely to the use of technology itself. The two distinct

pedagogical approaches that are cited most frequently in the research studies were the traditional teaching methods and the constructivist strategies. For example, Shyu (1999) investigated the effects of computer-assisted video-based anchored instruction of promoting students' attitudes toward mathematics and problem-solving skills. Focusing on 87 grade six Taiwanese students, she examined the effects of different media attributes on student mathematics achievement and attitudes in a situated learning environment. The results showed that anchored instruction enhanced student problems-solving skills. Students' attitudes, however, were not impacted by the technology supported anchored instruction.

In another study, Connell (1998) explored mathematics teaching and learning with technology in two rural classrooms during a one-year period. Both classrooms used technology but with different teaching approaches. One classroom adapted constructivist pedagogy and technology was used as a student tool for mathematics exploration. A behaviorist approach was used in the other class where technology was mainly used as a presentation tool. The results of the study confirmed a positive effect of technology on mathematics learning. By the end of each study both classrooms easily surpassed both state and district goals and had shown significant improvement from their baseline. Most importantly, however, the performance of the students in the constructivist class was significantly and consistently higher than that of the students in the other class. Further, the significant time by treatment effect suggested that the longer in which technology was used in this fashion, the greater the impact.

Xin (1999) examined the effects of combining CAI and a cooperative learning strategy. It was found that although math skills learning had increased for students using CAI in both cooperative and whole-class groups, there was a significant difference between the two on the post-test. One conclusion drawn from the findings was that math performance could be enhanced if students were given opportunities to work within a technology-assisted cooperative learning environment. Some researchers have noticed a shift in teaching and learning activities in the classroom as technology was integrated with the curriculum (Farrell, 1996; Ysseldyke, Spicuzza, Kosciolek, & Boys, 2003). Thus, when a positive outcome was claimed, it was difficult to determine if the gain was attributed solely to the technology intervention, to a particular instructional structure or to a combination of both. In their review paper, Christmann, Badgett, and Lucking (1997) claimed that many instructional factors, including cooperative learning, higher-order questions, and individualized instruction, can positively affect student learning outcomes.

Other factors that may contribute to mathematics teaching and learning with technology are the learner characteristics. Student achievement, their attitudes toward mathematics and technology, as well as their behaviors may depend on their gender, grade level, ability level, and their socio-economical status (SES).

Braden, Shaw, and Grecko (1991) evaluated a computer-assisted instructional (CAI) program for elementary hearing-impaired students in Florida. The results indicated that the CAI treatment had led to better in-class math quiz scores. Other outcomes such as reading and math scores on the Florida Statewide Student Achievement Tests (SSAT) were also measured, but no significant relationship was found.

Irish (2002) studied the effectiveness of a multimedia software program to teach students with learning and cognitive disabilities. Using a single-subject, multiple-baseline design across subjects (Cooper, Heron, & Heward, 1987), the study was systematically replicated across three pairs of grade 5 students. Although sample size (e.g. 6 students in total) was minimal, the results of this study showed that CAI could be an effective mechanism for teaching these special needs students certain mnemonic strategies, which in turn, could help increase their performance and accuracy on basic multiplication tasks.

In summary, the research reviewed on learning of mathematics with technology suggested that the effectiveness of mathematics learning with technology is highly depended on many other characteristics such as

teaching approaches, type of programs, and type of learners. Therefore the suggested study was included in this attempt of identifying the moderator study features used in this meta-analysis.

Methods

This meta-analysis quantitatively integrates the findings from the primary research on the educational use of technology in the teaching and learning of mathematics. The following section outlines the procedures employed to conduct this quantitative analysis.

Identification of Studies

In this study, the authors have focused on current researches, i.e. studies published in and after 1990. The rationale for choosing this time-frame is two-fold. First, high quality of technology use starts to surge in educational settings since this time because of a widespread appearance of microcomputers with its ever-increasing power, capabilities and lower prices in late 80's (Lau et al., 2001). Second, the existing meta-analysis about mathematics and technology has focused on studies conducted before 1990.

The study conducted a comprehensive search of the literature to locate appropriate studies. First, key topic-related descriptors as independent words were used to ensure a broad search of several computerized databases (see Dusek & Joseph, 1983). The initial step included an electronic search on the following databases: (a) Educational Resources Information Center (ERIC, 1990-2005), (b) PsycINFO (1990-2005), (c) Education Full Text (1995-2005), this database only have research papers published in the year of 1995 and after). Depending on the database, the search strategy varied and search terms included: *mathematics*, *math**, *mathematics learning*, *mathematics teaching*, and any term related to technology such as *technology**, *calculator**, or *Educational technology*.

The next step was to find, on the basis of the same descriptors, both qualitative and quantitative reviews published since 1990, as a means to enrich the pool of studies. Reference lists from, for example, Christmann, Badgett, and Lucking (1997), Clements (1998) and Woodward (1995) was checked for relevant studies. Finally, the study conducted a manual search of leading journals, namely: *Educational Technology Research and Development, British Journal of Educational Technology, Journal of Research on Technology in Education, and Journal of Computers in Mathematics and Science Teaching*, from the year1990 through the present.

Branching from primary studies and review articles, further appropriate citations were also identified. In this meta-analysis, every study had to meet the following inclusion/exclusion criteria:

- The study had to involve situations where K-12 students use technology.
- The study had to have data related to K-12 students' mathematics learning.
- The study had to report cognitive outcomes, behavior measures, and or affective outcomes. Different types of outcomes were coded and analyzed separately which are discussed in the section "outcomes and study features coding"; (For the the types of outcomes coded and analyzed; some outcomes were dropped due to small sample sizes.)
- The study had to report data that it was possible to examine, through the computation of effect sizes, the effects of technology on mathematics learning.

Using these criteria, abstracts from electronic searches, references from primary studies and review articles were examined to identify potential studies. Due to the limited resources available for this review, only English-language publications and databases were used. Approximately 500 abstracts were retrieved and reviewed. When in doubt, the study was collected and then read independently by the researchers and three graduate students for possible inclusion. A total of 138 papers were retrieved and reviewed for the study.

Outcomes and Study Features Coding

To identify methodological and substantive characteristics that might contribute to significant variations in the findings, outcomes and study features were coded using a three-stage coding procedure. First, a set of broad categories was established based on the

review of the related literature. The study features included four categories, namely publication features, sample descriptor, setting characteristics, and design characteristics. Further, outcome and methodological features were included in the coding scheme. Next, a random sample of 30% of the primary studies was selected. These studies were nomologically coded based on these categories to identify salient study features. Finally, the original coding scheme was revised and a codebook was created. Table 1 details the outcomes coded and Table 2 describes the study features coded.

Criteria of determination of independent findings

Many studies report results on multiple outcome measures (i.e. standardized mathematics achievement scores, school grades, attitudes) and in several occasions, two or more studies were reported in one paper (e.g. study 1 focused on regular students and study 2 examined learners who need special attention). Effect sizes for each measure were calculated and coded in the analysis.

Number of Findings Extracted

Two approaches, namely a single finding per study or multiple findings per study, were often used in meta-analysis regarding the number of findings to be extracted from each study. According to Lou et al. (2001), the advantage of extracting one finding per study was a guarantee of the independence of each finding. The disadvantage, however, was that the differences within a study between different sample groups (e.g. elementary vs. secondary students), or between different treatments under investigation (e.g. groups using one kind of computer system vs. another kind) were lost.

Extracting multiple effect sizes from a single study, however, might result in a violation of the independent

Outcome	Description
Achievement	Achievement scores measured by standardized or other tests.
Attitudes	Including students' attitudes (e.g. toward technology, calculators, mathematics).
Mathematics Behaviors	Including student behaviors such as problem solving, rote memorizing.

Table 1. Outcomes Coded

Study Features	Description
	Publication features
Publication Type	Was the study published in referred journals or unpublished proceedings/documents?
Publication year	Was the study reported in the last five years (i.e. 2000 and after) or earlier?
	Sample descriptor
SES	What was users socio-economic status (SES)? Was it low, middle, high, or mixed?
Age	Were the users elementary or secondary?
Race	What was the predominant race?
Gender	What was the gender composition? Was there less than 45% males (called 'female group'), or less than 45% females (called 'Male group'), or males and females were almost equal (called 'Mixed Group')?
Country	Where was the study conducted? Was it in North America, Europe, Australia, or other countries?
Student Type	Were they normal students or students with special needs (including at-risk, low achieving)
Type of Technology	What type of technology was considered? Was it mathematics software, or calculators or others (including the Internet, Spreadsheet)?
	Setting characteristics
Teaching Method	What was the teaching approach? Was it constructivist approach or traditional approach?
Length	How long did technology being used? Was it used for less than 16 weeks or longer?
	Design characteristics
Achievement instrument	Was the instrument non-standardized (school grades, teacher-or researcher- made tests), or standardized achievement scores?
Research design	Was it an experimental, quasi-experimental, or naturalistic study?

Table 2. Study Features Coded

assumption for effect sizes, which in turn, might increase Type I or II errors (Glass, McGaw, & Smith, 1981). In this study, two approaches were employed to resolve the dependence problem. "First, findings for each outcome were analyzed separately. Only one finding per outcome was extracted from each study unless they represented different subjects. This approach enabled one to examine different outcomes while ensuring independence among the findings for each outcome. Secondly, multiple effect sizes provided by the same subjects for the same category of outcome were dealt

with by randomly taking a single value from the set of correlated effect sizes per feature for each affected study. This methad eliminated the problem af dependency while ensuring that all levels of a study feature were represented" (Lou et al., 2001). In this study, grade levels (in crass-sectianal designs), and different ability graups (e.g. narmal students vs. special needs students) in a single study were considered separate primary studies (L. V. Hedges, 1987, personal cammunicatian, cited in Hyde, Fennema, & Laman, 1990).

Meta-analytic methadalagy literature is nat explicit on the use of longitudinal studies. Longitudinal data can be viewed as a single study in which correlations are aggregated ta represent the effect size af the study. Willett and Singer (1991) argued, however, that "a complex longitudinal time-dependent process cannot be adequately summarized by a single statistic" (p. 430). In line with this argument, langitudinal data in a study were treated as several independent primary studies based on different grade levels.

All the study findings were first extracted by the primary researcher. After this initial cading, a randamly selected 60% af the useful papers were recaded by a graduate assistant independently to test for reliability. The initial cading agreement an the number af findings to extract per study was 95.36%. Disagreements were resalved through further discussion and review of the disputed findings.

Effect Size Calculations

The effect size was calculated by the difference in the treatment and control group means divided by the paaled standard deviation (PSD). That is, the effect size was a measure of the effect of technology. Using the PSD is due to the homogeneity of variance in the population, in which case the PSD was mare stable and provides a better estimation af the papulatian variance than the control group SD alone (Hedges & Olkin, 1985; Hunter & Schmidt, 1990). Further, estimated effect sizes based on incamplete results (e.g. t- values, F- values, ANOVA tables, or p levels) were more readily comparable to effect sizes calculated in PSD (Lou et al., 2001).

Some studies did not report mean and SD, but provided data in the form of t-values, F-values, p-levels, frequencies, and/ar prapartians. The effect sizes af these studies were calculated using farmulas pravided by Lipsey and Wilson (2001). These estimations were computed or estimated using the software Effect Size Determinatian Pragram (Wilsan, 2001).

Same studies provided data callected from a different time periad (e.g. at the beginning ar the end af a semester). The correlation value was assumed as 0.70, when gain scores were available but pre-post correlatian was unavailable. This was based an the previaus experience with pre-and post-test results and two statisticians' (who do meta-analysis research) recommendatians (they cansidered that 0.70 was a canservative assumptian). There were a tatal af three studies in this category. Some studies only had post-tests in which case the post-test mean difference was the numeratar and the past-test PSD was the denaminatar. Some studies employed naturalistic design in which case there was no control group. In this case, a two-step appraach was used. First, the effect size was camputed cansidering pre-test scares as the cantral graup data and post-test scores as the treatment group data. This initial effect size was then multiplied by 2 which was the final estimated effect size af the study. Accarding ta Rasenthal (1991), "when the effect size estimates are the mean differences divided either by S or by (sigma), the definition af the size af the study changes by a factar af 2 in gaing from t for independent observatians ta t for carrelated observations" (p. 17).

The raw data for each finding were extracted separately by the researchers and a graduate student. After all the data were entered, the reliability was tested. The initial agreement was 89%. Disagreements were discussed and the conflict study findings were further reviewed until an agreement was reached.

Data Analysis

For each outcome, the unit of analysis was the independent study finding. Data screening was first

performed using the SPSS (SPSS, 2005) frequency and descriptive procedures. Several study features with almost no voriability (e.g. research design) or with over 90% missing data (e.g. subject leorned, tosk type, predominant race) were eliminated from further analysis. Categories within some variables (e.g. "ottitudes toward technology' ond 'ottitudes toward math') were collapsed bosed on the frequency distributions ond conceptual meaning.

Outlier analyses were performed by eliminating extreme volues from the effect size distribution (Lipsey & Wilson, 2001). Based on Lipsey and Wilson's (2001) recommendation, it was decided to exclude the studies that had effect size greoter than 3 standard deviotions from the meon of all the effect sizes.

The study tested the homogeneity of all effect sizes extracted from studies (Hedges & Olkin, 1985). First, effect sizes were corrected for bias and weighted by the inverse of its sampling voriance. Thot is, findings based on larger sample sizes were given more weight. The weighted effect sizes were then aggregated to form an overall weighted mean estimate of the effects (d_+). Homogeneity statistics (Q_T) was used to determine whether the set of effect sizes shared o common population parameter.

When the homogeneity statistics are significant, which signifies a heterogeneous set of effect sizes, two approoches can be used to ochieve the desired homogeneity. The first approoch is to delete outliers repeatedly until the remaining effect sizes become homogeneous. If this opproach foils to work, the second approach is that effect sizes ore divided into homogeneous subgroups (Hembree & Dessort, 1986).

Multiple Regression Model Tests

After the homogeneity tests, multiple regression models were tested using SPSS for Windows. When effect sizes were homogeneous, the population parameter could be determined by predictor voriobles. The substantive rationale is that studies differ because of different research design characteristics (Hedges & Olkin, 1985). One way was to fit homogeneous effect sizes into a general linear regression model throughout which we examined the effects of a number of independent variables on the dependent measure. The study employed the weighted least squares procedures for fitting general linear models as outlined in Hedges and Olkin (1985). Sample size were used to create weight for the regression onalysis (Hedges & Olkin, 1985) but not entered into the regression equation (Schrom, 1996).

Two weighted least squores multiple regression onalyses were performed for each outcome. Analysis one identified study features thot occounted for significant unique voriances in the findings. The significant predictors identified in onalysis I was then analyzed using the hierarchical weighted leost squares regressions (Hedges & Olkin, 1985) so that a parsimonious model could be developed (Lou et al., 2001). To better illustrate the overall effect (population coefficient) in this meta-analysis, the percentage of distribution non-overlap, or the U_3 statistic, was used to denote the chonge in scores or percentiles when a participant moved from one group to the other.

Results

In total, 81 independent effect sizes were extrocted from 39 studies involving a total of 59,147 students comparing mathematics learning with the use of technology versus mothematics leorning without technology on student achievement, attitudes, and behaviors. Appendix I provides details on each independent sample, including the number of students, defining chorocteristics of the independent sample, and the effect sizes calculoted. Close to half of the achievement outcomes were measured by non-standardized tests, mostly locally developed or teocher-made instruments or criteria specific to whot had been leorned on the computer tosks. Another 35 percent of the achievement outcomes were measured by standordized tests. Over sixty percent of the studies were well controlled, using either random assignment of students to experimental or control conditions or using statisticol control for quasi-experimental studies. Close to ninety percent of the papers were published in journal articles and about 10 percent were unpublished reports or conference proceedings.

Overall Effects

The homogeneity test of the 68 effect sizes of the achievement outcome was first conducted. It was significant which indicated that effect sizes were heterogeneous. Some evident outliers were deleted first in an attempt to improve the homogeneity of the remaining effect sizes. This approach, however, did not work. Another possible method was to divide the effect sizes into subgroups in order to identity homogeneous groups. The study adapted this approach and effect sizes associated with the achievement outcome were divided into 5 populations. One outlier, -2.41 in Mac Iver, Bonfanz, Plank (1998), was removed from the group with the smallest population parameter. The homogeneity tests were not significant at the .05 level for all the five populations (Q_1 = 20.09, df = 20; Q_2= 33.33, df = 25; Q_3 = 17.92, df = 11; Q_4= 1.53, df = 3; Q_5 = 1.90, df = 3) as detailed in Table 3.

The test of homogeneity of the attitude outcome was also significant which indicated that the effect sizes were heterogeneous. Two evident outliers, 6.82 in Shyu (2001) and 2.28 in Reimer & Moyer (2005), were deleted. The homogeneity test for the remain effect sizes was not significant at the .05 level (Q = 7.74, df=5).

Similarly, the homogeneity test for the behavior outcome was significant. After removing the outlier, -0.88 in Merriweather & Tharp (1999), the homogeneity test for the effect sizes was insignificant at the .05 level (Q= 7.27, df =3). Table 3 provided details for the three outcomes.

Moderator Analysis

According to Ma (1999), as much as the effect sizes all shared the same population difference, variation among effect sizes existed mainly because studies differed according to a number of research design characteristics. General linear regression was used to model this variation. This allowed us to identity the significant variables responsible for the variation among effect sizes and to gain insight into several practical concerns, such as gender differences and age differences. For the population parameters with 5 or less cases, statistical analysis were not conducted due to the

Outcome	k	Minimum	Maximum	d_+	Q	df
Achievement	68 (35)					
1st population parameter	21	-0.41	0.30	0.09	20.09	20
2nd population parameter	26	0.35	1.06	0.60	33.33	25
3rd population parameter	12	1.08	2.02	1.51	17.92	11
4th population parameter	4	2.04	2.84	2.40	1.53	3
5th population parameter	4	3.08	3.76	3.43	1.90	3
Attitudes	8 (8)					
		0.02	0.56	0.15	7.74	5
Behavior	5 (4)					
		0.52	1.12	0.86	7.27	3

Note: k is the total number of independent findings initially tested. The values in parentheses are the numbers of studies from which the findings were extracted. d_+ is the weighted mean effect size. Q is the homogeneity statistics. All the homogeneity statistics are not significant after the homogeneity procedure. df is the degree of freedom for the Q test.

Table 3. Population Parameters of the Effects of Technology on Schooling Outcomes in Mathematics Education

inadequacy of cases. Therefore, the regression and residual analysis were conducted only on the first three population parameters for the effects of technology on mathematics achievement. No further analysis was performed for the remaining population parameters. Table 4 displays the models considered in this meta-analysis.

In the effects of gender analysis, dummy coding was used to represent all the one-vector variables. That is, for the variables that comprised only two groups, they were recoded into dummy variables. These variables included: *race, grade level, country, treatment duration, technology type, instructional method, study type, achievement measure, publication year, and publication type*.

Effects of Gender

First population parameter:

Dummy coding was used to create two variables. The mixed group was used as the baseline against which the group of males less than 55% and the group of males more than 55% were compared. The results showed that the Q statistic (the weighted sums of squares) for gender grouping effects explained a statistically significant

Model	Regression test Q_R	df	Residual test Q_E	df
First population parameter (μ = 0.09, N = 21)				
Gender (2 vectors, male ≤ 55% vs. mixed, male > 55% vs. mixed)	114.41*	2	77.31*	18
Race (1 vector, mixed vs. non -mixed)	0.97	1	190.76	19
Socio-economic status (SES) (2 vectors, low vs. normal, mixed vs. normal)	9.16	2	182.57*	18
Grade level (1 vector, secondary vs. elementary)	109.17*	1	82.55*	19
Duration of treatment (1 vector, ≤ 1 year vs. > 1 year)	54.42*	1	137.31*	18
Type of technology (1 vector, software vs. others)	92.74*	1	98.99*	19
Achievement measure (1 vector, standardized vs. non-standardized)	1.66	1	190.07*	19
Year of publication (1 vector,≥ 2000 vs. < 2000)	1.21	1	190.52*	19
Type of publication (1 vector, journal article vs. others)	48.88*	1	142.85*	19
Gender, grade level	118.11*	3	73.62*	17
Gender, grade level, duration of treatment, type of technology, type of publication	120.92*	6	70.80*	14
Second population parameter (μ = 0.60, N = 26)				
Gender (2 vectors, male ≤ 55% vs. mixed, male > 55% vs. mixed)	106.63*	2	181.44*	23
Race (1 vector, mixed vs. non -mixed)	71.57*	1	216.51*	24
Socio-economic status (SES) (2 vectors, low vs. normal, mixed vs. normal)	31.92	2	256.15*	23
Grade level (1 vector, secondary vs. elementary)	57.08*	1	230.99	24
Country (1 vector, developing vs. developed)	47.06*	1	241.01*	24
Design (2 vectors, true experiment vs. natural, quasi experiment vs. natural)	38.29	2	249.88*	23
Duration of treatment (1 vector, ≤ 1 year vs. > 1 year)	37.54	1	250.53*	24
Type of technology (1 vector, software vs. others)	30.20	1	257.88*	24
Instructional method (1 vector, constructivist vs. traditional)	11.28	1	276.79*	24
Student type (1 vector, normal vs. at risk)	0.10	1	287.98*	24
Achievement measure (1 vector, standardized vs. non -standardized)	32.85	1	255.22*	24
Year of publication (1 vector, ≥ 2000 vs. < 2000)	89.45*	1	198.62*	24
Type of publication (1 vector, journal article vs. others)	2.28	1	285.79*	24
Gender, grade level	147.00*	3	141.07*	22
Gender, race, grade level, country, year of publication	170.94*	5	117.14*	20
Third population parameter (μ = 1.51, N = 12)				
Gender (2 vectors, male ≤ 55% vs. mixed, male > 55% vs. mixed)	11.00	2	71.28*	9
Race (1 vector, mixed vs. non -mixed)	0.80	1	81.47*	10
Socio-economic status (SES) (2 vectors, low vs. normal, mixed vs. normal)	16.42	2	65.86*	9
Grade level (1 vector, secondary vs. elementary)	3.21	1	79.07*	10
Country (1 vector, developing vs. developed)	0.80	1	81.47*	10
Design (2 vectors, true experiment vs. natural, quasi experiment vs. natural)	2.82	2	79.46*	9
Duration of treatment (1 vector, ≤ 1 year vs. > 1 year)	0.32	1	81.95*	10
Type of technology (1 vector, software vs. others)	2.44	1	79.84*	10
Instructional method (1 vector, constructivist vs. traditional)	5.93	1	76.35*	10
Student type (1 vector, normal vs. at risk)	2.80	1	79.48*	10
Achievement measure (1 vector, standardized vs. Non-standardized	6.78	1	75.50*	10
Year of publication (1 vector, ≥ 2000 vs. < 200)	2.07	1	80.21*	10

Note. * $p < 0.05$. The remaining two population parameters for the effects of technology on mathematics achievement, the population parameter for the effects of technology on attitude toward mathematics, and the population parameter for the effects of technology on mathematical behaviors do not have a sufficient number of studies for regression and residual analysis.

Table 4. Results of General Linear Regression Analysis of the Effects of Technology on Mathematics Achievement

amount of the variability in the effect sizes (Q_R = 114.41, df=2). However, the remaining variance, the Q statistic for error, was still statistically significant (Q_E = 77.31, df=18).

Second population parameter:

The analysis of the second population parameter showed similar results. Although the Q statistic for error was still statistically significant (Q_E = 181.44, df=23), the Q statistic for gender grouping effects explained a statistically significant amount of the variability in the effect sizes (Q_R = 106.63, df=2). Findings from both populations showed that gender grouping affected effect sizes of technology on achievement outcome when population parameters are small (0.09) and moderate (0.60).

Third population parameter:

For the third population parameter, the Q statistic for gender effects explained little (a statistically non significant amount) of the variability in the effect sizes (Q_R = 11.00, df=2). This finding showed that gender grouping had no impact on effect sizes of technology on achievement outcome when population parameter is large (1.51).

Effects of Race

The variable race was initially coded into 3 groups: grater than 60% white, greater than 60% minority, mixed with none more than 60%. Due to limited number of cases in certain groups, this was then collapsed into two categories: mixed vs. non-mixed (all one race).

First & third population parameter:

For both the first and third population parameters, the results showed that race explained a statistically non-significant amount of the variability in the effect sizes (Q_R = .97, df=1; Q_R = .80, df=1). This finding showed that race grouping did not affect effect sizes of technology on achievement outcome when population parameter is small (0.09) or big (1.51).

Second population parameter:

The analysis of the second population parameter showed a different result. Although the Q statistic for error was still statistically significant (Q_E = 216.51, df=24), the Q statistic for race effects explained a statistically significant amount of the variability in the effect sizes (Q_R = 71.57, df=1). This finding showed that race grouping affected effect sizes of technology on achievement outcome when population parameter is moderate (0.60).

Effects of Socio-Economic Status (SES)

Three socio-economic status groups were formed in this meta-analysis: low, normal, and mixed SES. Effect sizes were dummy coded so that the difference in low SES group was the baseline against which the differences in normal SES and mixed SES, respectively, were compared. For all three population parameters, the SES effects accounted for a small and statistically non-significant amount of the total variance (Q_R = 9.16, df=2; Q_R = 31.92, df=2; Q_R = 16.42, df=2 respectively). This finding showed that SES grouping did not affect effect sizes of technology on achievement outcome no matter how large population parameters are.

Effects of Grade Level

The variable grade level was initially coded as a continuous variable. Due to limited number of cases in certain groups, this was then collapsed into two categories: elementary and secondary.

First population parameters:

The results for the first population parameters showed that grade level explained a statistically significant amount of the variability in the effect sizes (Q_R = 109.17, df=1), even though the significant Q statistics for error (Q_E= 82.55, df = 19)indicated that grade level left out a significant amount of variance.

Second population parameter:

The analysis of the second population parameter showed a similar result. The Q statistic for grade level effects was a statistically significant (Q_R = 57.08, df=1), even though the Q value for error was statistically significant (Q_E = 230.99, df=24). Findings from both populations showed that grade level affected effect sizes of technology on achievement outcome when population parameters are small (0.09) and moderate (0.60).

Third population parameter:

Grade level had no significant effect on effect sizes of

technology on achievement outcome when populotion is large (1.51) (Q_R = 3.21, df=1).

Effects of Country

The variable grade level was coded in two groups: developing country and developed country. No statisticol analysis was performed for the first populotion porameter due to insufficient variance in the data.

Second population parameter:

The Q value for this population parameter showed a statistically significant result as country effects explained a small, though statistically significant amount of the total variance in the effect sizes (Q_R = 47.06, df=1). The Q value for error wos also statistically significant (Q_E = 241.01, df=24). Nevertheless, effect sizes were significant differently in different countries when population porometer is moderote (0.60).

Third population parameter:

In this populotion parometer, the Q stotistic for grade level effects explained little (o statistically non-significant amount) of the voriability in the effect sizes (Q_R = 0.80, df=1). This finding showed that countries shared similar effect sizes of technology on achievement outcome when population parameter is large (1.51).

Effects of Research Design

The voriable research design comprised three groups: true experiment, noturol, quasi experiment. Using true experiment as the baseline effect, the study effect coding to create two voriobles (see toble 4).

Because the doto did not have variance in research design, no analysis wos performed for the first populotion porameter. The analysis of the second ond third population parameters showed similar results. The Q statistic for research design was a not statisticolly significant (Q_R = 38.29, df=2; Q_R = 2.82, df=2 respectively). This finding showed that effect sizes did not vory significantly across research design when populotion effect sizes are moderate (0.60) and lorge (1.51).

Effects of Treatment Duration

Treatment duration was initially coded in three groups. Due to limited cases in certain groups, it wos collapsed into two groups: less or equal to one yeor vs. greater than one year.

First population parameter:

The results showed that treatment duration explained a stotistically significant omount of the variability in the effect sizes (Q_R = 54.42, df=1). The significant Q statistics for error (Q_E = 137.31, df=18), however, indicoted thot treatment duration left out a significant portion of the total voriance. Nevertheless, treatment duration was responsible for voriation in effect sizes of technology on achievement outcome when population parameter is smoll (0.09).

Second & third population parameter:

The analysis of the second and third population parameters showed similar results. The Q stotistic for treotment duration effects wos not stotisticolly significant (Q_R = 37.54, df=1; Q_R = 0.32, df=1). This finding showed that effect sizes did not vary significantly across treatment duration when population porometer is moderate (0.60) and large (1.51).

Effects of Technology Type

Technology type comprised two groups: software vs. other.

First population parameter:

Although the significont Q stotistics for error (Q_E = 98.99, df=19) indicated that technology type left out a significant portion of the total variance, technology type explained a statistically significant omount of the voriobility in the effect sizes (Q_R = 92.74, df=1). This finding showed that effect sizes were related significantly with technology type when population parameter is small (0.09).

Second & third population parameter:

The analysis of the second and third population parameters showed similar results. The Q stotistic for technology type effects was not stotisticolly significant (Q_R = 30.20, df=1; Q_R = 2.44, df=1 respectively). This finding showed thot effect sizes were not reloted significantly with technology type when populotion parameters are moderate (0.60) and large (1.51).

Effects of Instructional Method

The variable instructional methods comprised two groups: traditional vs. constructivist methods. Because there was not enough variance in the data for the first population parameter, no statistic analysis was performed for this parameter.

The analysis of the second and third population parameters showed similar results. The Q statistic for student type effects was not statistically significant (Q_R = 11.28, df=1; Q_R = 5.93, df=1 respectively). This finding showed that effect sizes did not vary significantly across instructional method when population parameters are moderate (0.60) and large (1.51).

Effects of Student Type

The variable student type comprised two groups: normal vs. special needs students. Using normal students as the baseline effect, effect coding was used to create two variables (Table 4). Due to the lack of variation in student type, data analysis was not conducted for the first population parameter. For the second & third population parameters, similar results have been found. The Q statistic for student type effects was not statistically significant (Q_R = .10, df=1; Q_R = 2.80, df=1 respectively). This finding showed that effect sizes did not vary significantly across student type.

Effects of Achievement Measure

Thinking that different measures of achievement might have effects, this variable was explored in two categories: standardized vs. non-standardized measure. The analysis of all three population parameters showed similar results: the Q values accounted for a statistically non-significant amount of the variability in the effect sizes (Q_R = 1.66, df=1; Q_R = 32.85, df=1 Q_R = 6.78, df=1, respectively). Effect sizes did not vary significantly across achievement measures.

Effects of Publication Year

The variable publication year was initially coded as a continuous variable. Due to limited number of cases in certain groups, this was then collapsed into two categories: before 2000 vs. year 2000 or later.

First & third population parameter:

For both first and third population parameters, similar results have been found: publication year explained a statistically non-significant amount of the variability in the effect sizes (Q_E = 1.21, df=1); (Q_E = 2.07, df=1). This finding showed that effect sizes had no relationship with when research studies were published when population parameters are small (0.09) and large (1.51).

Second population parameter:

Although the Q value for error was statistically significant (Q_E = 198.62, df=24), the Q value for publication year showed a statistically significant result which indicated that publication year explained a statistically significant amount of the total variance in the effect sizes (Q_R = 89.45, df=1). This finding showed that effect sizes depended on when research studies were published and when population parameter is moderate (0.60).

Effects of Publication Type

The variable publication type comprised two groups: journal articles vs. other.

First population parameter:

Although there is a significant Q statistic for error (Q_E = 142.85, df=19), publication type explained a statistically significant amount of the variability in the effect sizes (Q_R = 48.88, df=1). This finding showed that effect sizes depended on where research studies were published when population parameter is small (0.09).

Second & third population parameter:

For the second population parameter, the Q statistic for publication type effects was not statistically significant (Q_R = 2.28, df=1). This finding showed that effect sizes did not vary significantly across publication type when population parameter is moderate (0.60). The analysis for the third population parameter was not performed due to the lack of variation in publication type.

The Final Model

The final model included all individually significant variables as discussed in individual models above. This analysis was performed for the first and second population parameters only. For the third population

porameter, since the Q statistic was not significant for all the voriables analyzed in the individuol models, no further analysis was performed. Table 5 provides details.

First population parameter:

Results for the first population porometer essentiolly emphasized two predictors: gender and grade level. The Q statistic for the model with treatment durotion, technology type, and publication type in oddition to gender and grade level indicated thot this model explained a statistically significant and practically substantiol amount of varionce in the effect sizes (Q_R = 120.92, df=6). However, it was noticed thot this model added little to the model with gender and grade level alone (Q_R = 118.11, df=3). This is a good indicotion that gender and grode level were the most important predictors of effect sizes of technology on achievement outcome when population parameter is small (0.09).

Second population parameter:

The same is true for the second population parameter. When populotion parometer is moderote (0.60), gender and grade level were the most importont predictors of effect sizes of technology on achievement outcome though to a lesser degree than the first population porameter (Q_R = 170.94, df=5 versus Q_R = 147.00, df=3).

All individually statistically significant voriobles were tested together not only for varionce estimation but also for relotive importance. Some individually significant variables became non-significant in the presence of other variables. These non-significant voriables were removed one by one starting from the one with lorgest *p* value. After these deletions, remaining variables were all statisticolly significant. These variables are relatively important predictors of effect sizes of technology on achievement outcome. Table 5 shows the effects. For the first populotion porometer, the most importont predictor was grade level, with secondary school students showing significantly larger effects than elementary school students. In other words, technology had bigger effects on secondory than elementory school students. The same technology applied at the secondory level would improve achievement of secondary students from 50th percentile to 54th percentile.

For the second populotion porometer, gender ond grade level were the most important predictors of effect sizes of technology on achievement outcome, with grade level showing slightly more important effects than gender. Gender group with less than or equol to 55% moles outperformed mixed gender group under the same technology. For a student, moving from the mixed gender group to the other gender group wos associated with an increase in ochievement outcome from the 50th to the 58th percentile. Grade level was reversed between the first and second population parameters. The same technology when applied at the elementary school level would improve achievement of elementary school students from 50th percentile to 62nd percentile.

Model	Effect	SE	U_3
1st population parameter			
Grade level (secondary vs. elementary)	0.13	0.03	54%
2 population parameter			
Gender (male ≤ 55% vs. mixed)	0.17	0.04	58%
Grade level (secondary vs. elementary)	-0.25	0.10	62%

*Note. All effects are statistically significant at the alpha level of 0.05.

Table 5. Statistically Significant Predictors for the Variation among Primary Studies in Terms of the Effects of Technology on Mathematics Achievement

Discussion

This study sheds light on the use of technology in school mathematics learning. Based on a total of 81independent findings extracted from 39 studies involving a total of 59,147 learners, the results of the series of meta-analyses conducted in this review indicate that technology can affect mathemotics teaching ond leorning. Although difficult to offer a comprehensive recipe for the design of technology-supported learning environment, it is possible to offer some practical guidelines for technology integrotion in mathematics leorning.

The most important finding of this study is that, in general, gender and grade level are the two most important predictors for the effect of technology on mathematics achievement outcome. Specifically, in the first population, the grade level significantly predicted the

impact of technology on mathematics achievement. The effects of technology were significantly enhanced when the same technology was used in secondary students rather than elementory students. In the second population, both grade level and gender significantly influenced mathemotics ochievement. Thot is, students' math ochievement is significantly improved with the use of technology when they are (i) in elementary schools, and (ii) in groups with greoter thon 55% females.

Regardless of sample populations, grade level is always a strong factor that contributes to the effect of technology on moth achievement. This suggests thot when planning technology integration projects, we need to consider elementary and secondary schools separately to ensure appropriate approaches for each grade level. It is interesting to see thot the impact of technology is reversed between the first and second population porometers. Further reseorch into this phenomenon is needed.

Another interesting finding is that technology can significantly improve math achievement outcomes when students are in female dominated groups rather than gender balanced groups. One possible explanation for this is that females feel more comfortoble and confident when they are in female dominated groups (Li, 2005; Mullany, 2000; Savicki, Kelley, & Lingenfelter, 1996b). Therefore, when implementing technology in schools, we need to consider grouping female students together.

Quality Assessment

It is importont to notice that, confidence roting for effect size colculation is generally high for the research studies used in this review. Over 65 percent of the effect sizes are computed with no estimation using the descriptive data such as means, stondard deviotions, frequencies, etc. Another 22 percent of the effect sizes are slightly estimated using significance testing statistics with complete statistics of conventional sort rather than descriptive statistics. Only about 12 % of the effect sizes are calculated with some estimation or highly estimated.

The study also evaluated the quality of research studies used in this review. Of all the studies, over 43 % of them used quasi-experimental design which involved experimental and control groups. Another 20% of the studies used not only treotment and control group design, but also employed randomized selection of samples of same sort. Only 37% of the research used naturalistic design. This, coupled with the foct that majority of the effect sizes were highly relioble, indicates that the studies used in this review are mainly of "high methodological quolity" (Lipsey & Wilson, 2001, p.157).

Strengths, Limitations, and Future Directions

This meta-analysis extends knowledge of the effect of technology on mothematics teaching and learning focusing on cognitive and affective outcomes. It has addressed the question of whether and to what extent mathematics learning with technology is more effective than without the use of technology and on which outcomes. It has identified some study features thot moderated the effects of technology on mathematics achievements. Through weighted least squores multiple regression anolyses, models were developed that accounted for the variability of technology effects on achievement outcomes.

This meta-analysis, however, just like any research, has its limitations. First, "as meta-analysts do not have experimental control of data" (Lou et al., 2001, p. 482), we have to use some studies with small somple sizes. Sometimes, we have to estimate the effect sizes when there ore missing data. This, inevitably, reduces the sensitivity of the anolysis. Second, multiple regression analyses are sensitive to the order variables are entered. The regression models were identified by the study, by no means, are finol and conclusive. Finally, the design quality of the programs used in the primary studies may offect the quality of this meta-analysis.

As technology has become ubiquitous in education, including mathemotics teaching and learning, we may leorn more about the empowering effects of technology. Finding better ways to integrate technology in education to meet the diverse needs of students is an ongoing task for educators and researchers.

Acknowledgment:

This reseorch was partially supported by o research grant from the American Institute of Research (AIR) to the authors. The views, findings and opinions expressed here are those of authors and do not necessarily reflect the views or positions of AIR.

References

[1]. **Albright, M., & Graf, D. (1992).** *Teaching in the Information Age: The Role of Educational Technology*. San Froncisco: Jossey-Bass Publishers.

[2]. **Ball, S. (1988)**. Computers, concrete materials ond teaching froctions. *School science and mathematics., 88*, 470-475.

[3]. **Braden, J. P., Shaw, S. R., & Grecko, L. (1991)**. An Evaluotion of a Computer-Assisted Instructional Progrom for Elementary Hearing-Impaired Students. *The Volta Review, 93*(6), 247-252.

[4]. **Brush, T. A. (1997)**. The effects of group composition on ochievement and time on task for students completing ILS activities in cooperative pairs. *Journal of Research on Computing in Education, 30*(1), 2-17.

[5]. **Char, C. (1989)**. *Computer graphics feltboards: New software approaches for young children's mathematical exploration*. San Francisco: Americon Educotionol Research Association.

[6]. **Christmann, E., Badgett, J., & Lucking, R. (1997)**. Microcomputer-Based Computer-Assisted Instruction within Differing Subject Areas: A Statisticol Deduction. *Journal of Educational Computing Research, 16*(3), 281-296.

[7]. **Clark, R. (1983).** Reconsidering reseorch on learning from media. *Review of educational research, 53*, 445-460.

[8]. **Clements, D. (1998)**. /Young Children ond Technology./ Paper presented at the Forum on early childhood science, mathemotics, and technology education, Woshington, DC.

[9]. **Coley, R., Cradleer, J., & Engel, P. (2000)**. *Computers and the Classroom: The Status of Technology in US Schools*. Princeton, NJ: Policy Information Center, Educational Testing Service.

[10]. **Connell, M. L. (1998).** Technology in Constructivist Mothematics Classrooms. *The Journal of Computers in Mathematics and Science Teaching, 17*(4), 311-338.

[11]. **Cooper, J., Heron, T., & Heward, W. (1987).** *Applied Behavior Analysis*. Columbus: Merrill.

[12]. **Dusek, J. B., & Joseph, G. (1983)**. The boses of teacher expectoncies: A meta-onolysis. *Journal of Educational Psychology, 75*, 327-346.

[13]. **Farrell, A. M. (1996).** Roles ond Behaviors in Technology-Integroted Precolculus Classrooms. *Journal of Mathematical Behavior, 15*, 35-53.

[14]. **Funkhouser, C. (1993).** The Influence of Problem-Solving Software on Student Attitudes obout Mothematics. *Journal of Research on Computing in Education, 25*(3), 339-346.

[15]. **Glass, G. V., McGaw, B., & Smith, M. L. (1981).** *Meta-analysis in social research*. Beverly Hills, CA: SAGE.

[16]. **Hecht, J. B., Roberts, N. K., Schoon, P. L., & Fansler, G. (1995)**. *Teacher Teams and Computer Technology.* Paper presented at the 1995 Annual Meeting of the Mid-Western Educational Reseorch Associotion, Chicago, IL.

[17]. **Hedges, L., & Olkin, I. (1985)**. *Statistical methods for meta-analyses*. Orlando, FL: Academic Press.

[18]. **Hembree, R., & Dessart, D. J. (1986)**. Effects of hond held colculotors in precollege mathemotics education: A meto-anolysis. /Journol for Research in Mathematics Education, 17/, 83-99.

[19]. **Hunter, E., & Schmidt, F. (1990).** *Methods of meta-analysis: correcting error and bias in research findings*. Newbury Park, CA: Sage.

[20]. **Hyde, J. S., Fennema, E., & Lamon, S. J. (1990)**. Gender differences in mathematics performonce: A meto-anolysis. *Psychological Bulletin, 107*, 139-155.

[21]. **Irish, C. (2002).** Using Peg- and Keyword Mnemonics and Computer-Assisted Instruction to Enhance Bosic Multiplicotion Performance in Elementary Students with Leorning and Cognitive Disabilities. *Journal of Special Education Technology, 17*(4), 29-40.

[22]. **Jones, K. (2005)**. Graphing calculators in the

teaching and learning of mathematics: A research bibliography. *Micromath, Summer*, 31-33.

[23]. **Kieran, C., & Hillel, J. (1990)**. It's tough when you have to make the triangles angles: Insights from a computer-based geometry environment. *Journal of Mathematical Behavior, 9*, 99-127.

[24]. **Kim, S. (1993)**. The relative effectiveness of hands-on and computer-simulated manipulative in teaching seriation, classification, geometric, and arithmetic concepts to kindergarten children. *Dissertation Abstracts International, 54*(09), 3319.

[25]. **Kulik, K. (2003)**. *Effects of using instructional technology in elementary and secondary schools: What controlled evaluation studies say* (Research Report No. SRI Project Number P10446.003). Arlington, VA: SRI International.

[26]. **Kulik, K., Schwalb, B., & Kulik, J. (1982)**. Programmed instruction in secondary education: A meta-analysis of evaluation findings. *Journal of Educational Research, 75*(3), 133-138.

[27]. **Li, Q. (2004)**. *Technology and mathematics education: Any impact?* Paper presented at the the Eleventh International Literacy and Education Research Network Conference on Learning, La Havana.

[28]. **Li, Q. (2005)**. Gender and CMC: A review on conflict and harassment. *Australasian Journal of Educational Technology, 21*(3), 382-406.

[29]. **Lipsey, M., & Wilson, D. (2001)**. *Practical meta-analysis* (Vol. 49). Thousand Oaks: SAGE Publications.

[30]. **Lou, Y., Abrami, P. C., & d'Apollonia, S. (2001)**. Small group and individual learning with technology: A meta-analysis. *Review of educational research, 71*(3), 449-521.

[31]. **Mac Iver, D. J., Balfanz, R., & Plank, S. B. (1998)**. The Talent Development Middle School. An Elective Replacement Approach to Providing Extra Help in Math - The CATAMA Program (Computer- and Team-Assisted Mathematics Acceleration). Report No. 21/. Baltimore, MD: Center for Research on the Education of Students Placed at Risk.

[32]. **Ma, X. (1999)**. Anxiety and Achievement. *Journal of Research in Mathematics Education*.

[33]. **Merriweather, M., & Tharp, M. L. (1999)**. The Effect of Instruction with Graphing Calculators on How General Mathematics Students Naturalistically Solve Algebraic Problems. *The Journal of Computers in Mathematics and Science Teaching, 18*(1), 7-22.

[34]. **Moyer, P. S., & Bolyard, J. (2002)**. Exploring representation in the middle grades: Investigations in geometry with virtual manipulatives. *The Australian Mathematics Teacher, 58*(1), 19-25.

[35]. **Moyer, P. S., Bolyard, J., & Spikell, M. (2002)**. What are virtual manipulatives? *Teaching Children Mathematics, 8*(6), 372-377.

[36]. **Mullany, L. (2000)**. The application of current language and gender theory to managerial meeting discourse. *Nottingham Linguistic Circular, 15*, 1-16.

[37]. **Mushi, S. (2000)**. *Use of Interactive Video Technology to Teach Middle School Mathematics in Chicago*. Chicago: Northeastern Illinois Univ., Chicago, Dept. of Teacher Education.

[38]. **Nute, N. (1997)**. The impact of engagement activity and manipulatives presentation on intermediate mathematics achievement, time-on-task, learning efficiency and attitude. *Dissertation Abstracts International, 58*(08), 2988.

[39]. **Parr, J. (2003)**. *A review of the literature on computer-assisted learning, particularly integrated learning systems and outcomes with respect to literacy and numeracy*. Auckland, New Zealand: M. O. Education. The University of Auckland.

[40]. **Reimer, K., & Moyer, P. S. (2005)**. Third-Graders Learn About Fractions Using Virtual Manipulatives: A Classroom Study. *Journal of Computers in Mathematics and Science Teaching, 24*(1), 5-25.

[41]. **Rosenthal, R. (1991)**. *Meta-analytic procedures for social research* (Vol. 6). Newbury Park, CA: SAGE Publications Ltd.

[42]. **Royer, J. M., Greene, B. A., & Anzalone, S. J. (1994)**. Can U.S. developed CAI work effectively in a developing country? *Journal of Educational Computing Research,*

10(1), 41-61.

[43]. **Salerno, C. A. (1995).** The Effect of Time on Computer-Assisted Instruction for At-Risk Students. *Journal af Research an Camputing in Educatian*, *28*(1), 85-97.

[44]. **Savicki, V., Kelley, M., & Lingenfelter, D. (1996b).** Gender and graup campasitian in small task graups using computer-mediated communication. *Computers in human behavior, 12*(2), 209-224.

[45]. **Schram, C. M. (1996).** A meta-analysis of gender differences in applied statistics achievement. *Journal of Educatianal and Behaviaral Statistics, 21*(55-70).

[46]. **Shyu, H. Y. (1999).** Effects af Media Attributes in Anchared Instruction. *Jaurnal af Educational Camputing Research, 21*(2), 119-139.

[47]. **Shyu, H. Y. (2000).** Using video-based anchored instruction to enhance learning: Taiwan's experience. *British Jaurnal af Educatianal Technalagy, 31*(1), 57-69.

[48]. **Suh, J., Moyer, P. S., & Heo, H. (2005).** Examining technalagy uses in the classraam: Develaping fractian sense using virtual manipulative concept tutorials. *The Journal of Interactive Online Learning, 3*(4), 1-21.

[49]. **Terry, M. (1996).** An investigation of differences in cagnitian when utilizing math manipulatives and math manipulative software. *Dissertation Abstracts International, 56*(7), 26-50.

[50]. **Thompson, P. (1992).** Nototions, conventions, ond canstraints: Contributians ta effective uses af cancrete materials in elementary mathematics. *Jaurnal af Research in Mathematics Education, 23*, 123-147.

[51]. **Willett, JB, & Singer, JD. (1991).** From whether to when: New methods for studying student dropout ond teacher attritian. /Review af Educatianal Research/, 61, 407-450.

[52]. **Wilson, D. (2001).** Effect Size Determinatian Pragram. Maryland: University of Maryland.

[53]. **Woodward, J. (1995).** Technolagy-Based Research in Mathemotics for Speciol Education. /Focus on Learning Prablems in Mathematics, 17/(2), 3-23.

[54]. **Xin, J. F. (1999).** Computer-Assisted Cooperative Learning in Integrated Classraams far Students With and Without Disabilities. *Information Technology in Childhood Education*, 61-78.

[55]. **Ysseldyke, J., Spicuzza, R., Kosciolek, S., & Boys, C. (2003).** Effects of a Learning Information System on Mathematics Achievement and Classraam Structure. *The Journal of Educational Research (Washington, D.C.), 96*(3), 163-173.

Appendix 1. Effect Sizes for Studies of Technology

Study	Grade-level	Outcome Extracted	Characteristics that Distinguish the Findings within the Study	N	ES
Xin (1999)	Grade 3	achievement	Narmal students, callabarative learning	46	1.92
---	Grade 3	achievement	Narmal students, whale class learning	47	.74
---	Grade 3	achievement	Special needs students, callabarative learning	13	2.84
---	Grade 3	achievement	Special needs students, callabarative learning	12	2.60
Farrell (1996)	High schaal	behaviar		180	.520
Flener (2000)	Grade 5-8	Attitude		408	.160
Shyu (1999)	Grade 6	achievement		53	.83
---	Grade 6	achievement		53	.09
Marena & Mayer (1999)	Grade 6	achievement	High achieving students, teacher made tests	46	1.34
---	Grade 6	achievement	Law achieving students, teacher made tests	26	3.32
---	Grade 6	achievement	Narmal students, camputer-based tests	60	3.08

Funkhauser (1993)	High schaal	attitude		40	.024
---	High school	achievement		71	.720
Hecht (1995)	Grade 9	achievement		104	.39
Martindale, Pearsan et al. (2005)	Grade 5-10	achievement	Grade 5 students taking math in 2001	876	.43
---	Grade 5-10	achievement	Grade 5 students taking math in 2002	970	.61
---	Grade 5-10	achievement	Grade 8 students taking math in 2001	1869	.06
---	Grade 5-10	achievement	Grade 8 students taking math in 2002	2094	.06
---	Grade 5-10	achievement	Grade 10 students taking math in 2001	2267	.03
---	Grade 5-10	achievement	Grade 10 students taking math in 2002	2499	.07
Shyu (2000)	Grade 5	attitudes		74	6.82
---	Grade 5	achievement		74	1.44
Ysseldyke, Spicuzza et al (2003)	Grade 4/5	achievement	Grade 4 students	6542	.19
---	Grade 4/5	achievement	Grade 5 students	6542	.35
---	Grade 4/5	behaviar	Grade 4/5 students	87	1.12
Ysseldyke, Tardrew, Betts et al (2004)	Grade 3-6	achievement	Gifted students	100	.45
---	Grade 3-6	achievement	Narmal students	1479	.47
Ysseldyke, Betts, Thill & Hannigan (2004)	Grade 3-6	achievement		270	.500
Iskander & Curtis (2005)	High schaal	achievement		43	2.02
Reimer & Mayer (2005)	Grade 3	achievement		19	.66
---	Grade 3	attitudes		19	2.28
Schpliberg & Hubschman (2003)	High schaal	achievement		56	.08
Wittman, et al (1998)	Grade 4	achievement		24	.83
Quinn & Quinn (2001a)	Grade 3-5	achievement		88	2.12
Quinn & Quinn (2001b)	Grade 3-5	achievement		77	2.04
Wang, Wang & Ye (2002)	High schaal	achievement		24199	.04
Feng & Josephine (2000)	Kindergarten students	achievement		47	.13
Braden, Shaw & Grecka (1991)	grade 1	achievement		48	.41
McBride & Lewis (1993)	k-12	Achievement		31	1.42
---		attitude		95	.56

Rayer, Greene & Anzalane (1994)	Elementary	Achievement	Grade 9 (1987) in schaal 1	173	-.075
---	Elementary	Achievement	Grade 9 (1987) in schaal 2	133	.811
---	Elementary	Achievement	Grade 9 (1987) in schaal 3	170	.082
-	Elementary	Achievement	Grade 9 (1987) in schaal 4	185	.215
---	Elementary	Achievement	Grade 9 (1987) in schaal 5	160	1.060
---	Elementary	Achievement	Grade 5 (1988) in schaal 1	177	.490
---	Elementary	Achievement	Grade 5 (1988) in schaal 2	128	.791
---	Elementary	Achievement	Grade 5 (1988) in schaal 3	165	1.080
---	Elementary	Achievement	Grade 5 (1988) in schaal 4	183	-.413
---	Elementary	Achievement	Grade 5 (1988) in schaal 5	127	.301
---	Elementary	Achievement	Grade 5 (1988) in schaal 1	173	.380
---	Elementary	Achievement	Grade 5 (1988) in schaal 2	189	.420
---	Elementary	Achievement	Grade 5 (1988) in schaal 3	127	.610
---	Elementary	Achievement	Grade 5 (1988) in schaal 4	156	.220
---	Elementary	Achievement	Grade 5 (1988) in schaal 5	187	1.750
Blantan, Maarman, Hayes, Warner (1997)	Grade 3-6	Achievement		52	1.43
Chute & Miksad (1997)	Kindergarten students	Achievement		51	.29
Wheeler & Regian (1999)	Grade 9	Achievement		493	.026
Ysseldyke,Spicuzza, Kascialek Teelucksingh, Bays, & Lemkuil (2003)	Elementary	Achievement	Students in schaal 1	881	.13
---	Elementary	Achievement	Students randamly selected fram district	826	.14
Clariana (1996)	Grade 5	Achievement		873	.63
Salerna (1995)	Grade 5	Achievement		119	1.604
Merriweather & Tharp (1999)	Grade 8	Attitude		80	.07
---	Grade 8	Behaviar		52	-.88
Olkun (2003)	Grade 4/5	Achievement		62	.67
Cannell (1998)	Elementary	Achievement	Students in graup 1	25	3.76
---	Elementary	Achievement	Students in graup 2	27	3.56
Mac Iver, Balfanz, & Plank (1998)	Grade 7	Achievement	High achiever students	88	.68
---	Grade 7	Achievement	Law achiever students	8	-2.41
Irish (2002)	Grade 4/5	Achievement		6	.26
Funkhauser (2003)	Grade 10/11	Achievement		49	.412
---	Grade 10/11	Attitude		49	.017

Page (2002)	Elementary	Achievement		207	.904
---	Elementary	Behavior	Students enrolled in 1998	207	.780
---	Elementary	Behavior	Students enrolled in 1999	207	1.020
Carter & Smith (2002)	High school	Achievement	Students in Algebra I class	228	.046
---	High school	Achievement	Students in Algebra II class	228	.080
Graham, Thomas (2000)	High school	Achievement	School A	94	1.180
---	High school	Achievement	School B	94	1.190
Kalyuga & Sweller (2005)	Grade 10	Achievement	Group 1 with learner-adapted format	30	1.720
---	Grade 10	Achievement	Group 2 with nonlearner-adapted format	30	.426

ABOUT THE AUTHORS

** Associate Professor, Faculty of Education , University of Calgary*
*** Professor, Department of Curriculum and Instruction, University of Kentucky*

Dr. Qing Li is an Associate Professor in the Faculty of Education at the University of Calgary. Dr. Li's research interest includes: educational technology, mathematics education, technology integration in math and science, equity, and cyberbullying.

Xin Ma is a Professor in the Department of Curriculum and Instruction at the University of Kentucky and a Fellow of the (U.S.) National Academy of Education. His areas of specialization include mathematics education, school effectiveness, policy research and quantitative methods in educational research and evaluation. He has been engaged in analyzing large-scale local, national, and international databases for the evaluation of educational policies, interventions, and innovations for many years.